Automatic pre-screening of outdoor airborne microplastics in micrographs using deep learning

Sheen Mclean Cabaneros [a] [iD],*, Emma Chapman [b], Mark Hansen [c] [iD], Ben Williams [d] [iD], Jeanette Rotchell [b,e]

[a] *School of Engineering, University of Hull, Kingston upon Hull, HU6 7RX, UK*
[b] *School of Natural Sciences, University of Hull, Kingston upon Hull, HU6 7RX, UK*
[c] *Centre for Machine Vision, School of Engineering, University of the West of England, Bristol, BS16 1QY, UK*
[d] *Air Quality Management Resource Centre, University of the West of England, Bristol, BS16 1QY, UK*
[e] *College of Health and Science, University of Lincoln, Lincoln, LN6 7TS, UK*

ARTICLE INFO

ABSTRACT

Airborne microplastics (AMPs) are prevalent in both indoor and outdoor environments, posing potential health risks to humans. Automating the process of identifying potential particles in micrographs can significantly enhance the research and monitoring of AMPs. Although deep learning has shown substantial promise in microplastics analysis, existing studies have primarily focused on high-resolution images of samples collected from marine and freshwater environments. In contrast, this work introduces a novel approach by employing enhanced U-Net models (Attention U-Net and Dynamic RU-NEXT) along with the Mask Region Convolutional Neural Network (Mask R-CNN) to identify and classify outdoor AMPs in low-resolution micrographs (256×256 pixels). A key innovation involves integrating classification directly within the U-Net-based segmentation frameworks, thereby streamlining the workflow and improving computational efficiency. This marks an advancement over previous work where segmentation and classification were performed separately. The enhanced U-Net models attained average classification F1-scores exceeding 85% and segmentation accuracy above 77% on test images. Additionally, the Mask R-CNN model achieved an average bounding box precision of 73.32%, a classification F1-score of 84.29%, and a mask precision of 71.31%. The proposed method provides a faster and more accurate means of identifying AMPs compared to thresholding techniques. It also functions effectively as a pre-screening tool, substantially reducing the number of particles requiring labour-intensive chemical analysis. By integrating advanced deep learning strategies into AMPs research, this study paves the way for more efficient monitoring and characterisation of microplastics.

1. Introduction

Microplastics (MPs), plastic particles of less than 5 mm in diameter, have gained global attention due to their widespread presence across marine, terrestrial, and atmospheric environments (Chen et al., 2020; Kim et al., 2020; Hale et al., 2020). MPs have been detected in various human clinical samples, including the lungs, blood, colon, and liver (Jenner et al., 2022; Leslie et al., 2022; Horvatits et al., 2022; Ibrahim et al., 2021). Airborne microplastics (AMPs) is an increasing environmental concern due to their abundance indoors and outdoors (Zhao et al., 2023; Netema et al., 2024). With people spending 90% of their time indoors, AMPs can enter the human body via inhalation and ingestion (Dewika et al., 2023; Vattanasit et al., 2023; Netema et al., 2024), and migrate to soil and marine environments via precipitation and natural deposition (Su et al., 2023). Since MPs are known to carry harmful pollutants in marine environments (Zarfl and Matthies, 2010), they likely transport pollutants through the air as well (Wright and Kelly, 2017). Consequently, some researchers even

argue that MPs in the air may pose greater health risks than those found in water, food, or soil (Pauly et al., 1998; Vattanasit et al., 2023).

Currently, no comprehensive protocol exists for monitoring and detecting AMPs. Various methods are employed for MPs collection (e.g., passive or active sampling), extraction (e.g., density separation), and identification (e.g., visual or instrumental analysis) (Shao et al., 2022). However, the results are often reported in incompatible units, complicating comparisons with breathable volume units used in statutory air quality monitoring (Jenner et al., 2022; Jahandari, 2023). As suggested by Hartmann et al. (2019), a standardised database categorising MPs based on size, shape, colour, and origin is needed to better understand their global impact.

An ideal pipeline for monitoring AMPs would include both chemical and physical identification techniques. Chemical characterisation methods such as Fourier-transform Infrared (FTIR) spectroscopy, Raman spectroscopy, energy-dispersive X-ray Spectrometry (EDS), and pyrolysis gas chromatography-mass spectrometry (Py-GC-MS) are commonly used (Su et al., 2023). On the other hand, physical characterisation methods such as stereoscopic and scanning electron microscopy (SEM) help capture images and analyse particle attributes (Shao et al., 2022). Characteristics such as size, levels, and shape can indicate the origin of MPs and potential toxicity outcomes in human cell/tissue experiments (Danopoulos et al., 2022). However, these methods are time-consuming and costly, limiting their potential for high throughput analyses of samples (Shao et al., 2022; Cui et al., 2023). Emerging automated techniques can help detect MPs in images, serving as an initial screening method that reduces the need for extensive chemical analysis (Primpke et al., 2020).

The identification of potential MPs in images involves segmentation (isolation of particles from the background), and classification (categorisation of objects by shape). Traditional methods include visual identification, boundary-based image processing (Mukhanov et al., 2019; Gauci et al., 2019), and thresholding techniques (Otsu, 1979; Sauvola and Pietikäinen, 2000). While visual inspection is simple and low-cost, it is prone to errors for particles smaller than 2 mm and cannot be easily automated (Kroon et al., 2018). Boundary-detection methods can be automated but depend heavily on image quality, and poor resolution or noise can reduce their accuracy (Bovik, 2009). Thresholding techniques are quicker to implement but are sensitive to image noise and rely on fixed thresholds, making them less adaptable to varying image conditions (Gonzales-Barron and Butler, 2006). Variations in particle characteristics such as diameter, circularity, and roughness across images from different devices further complicate the standardisation of MPs databases (Shi et al., 2022).

Deep learning techniques based on Convolutional Neural Networks (CNN) have been increasingly adopted to identify MPs in images in recent years (Lin et al., 2022; Su et al., 2023; Zhu et al., 2023). One prevalent approach employs U-Net and VGG16 neural networks for image segmentation and classification. For instance, U-Net has been utilised to segment images of MPs from beach samples captured by digital cameras (Lorenzo-Navarro et al., 2021) and SEMs (Shi et al., 2022), while VGG16 has been applied to classify the shapes of detected MPs. Although these studies have achieved impressive results, segmentation and classification tasks are performed separately. Another approach involves the use of the Mask Region CNN (Mask R-CNN) model which allows simultaneous segmentation and classification of particles (He et al., 2016). For instance, Mask R-CNN has been employed to segment and classify MP fibres in images (Wegmayr et al., 2020) and to analyse beach sediments captured by digital cameras (Han et al., 2023) and Zeiss stereo microscopes (Huang et al., 2023).

Current research on deep learning for MPs identification predominantly focuses on samples from marine, freshwater, or ocean environments, where particles often originate from runoff, wastewater, or direct disposal. These environments present unique challenges, such as interference from organic matter, salts, variable flow rates, and tidal movements, necessitating tailored approaches for accurate MPs identification. Additionally, previous studies have used high-resolution images (at least 512×512 pixels) with varying particle sizes, background ratios, and intensities to train deep learning models. Dealing with AMPs presents significant challenges as they are sparsely distributed, unlike MPs in water or sediment environments (Adhikari et al., 2022). This limited sample availability results in fewer images for analysis, providing fewer reference features for deep learning models to extract. Additionally, AMPs in micrographs may appear transparent or semi-transparent following treatments designed to remove background particles by digesting organic material (Chapman et al., 2024). Consequently, segmentation and classification become more challenging and prone to errors. Hence, this work broadens the application range of deep learning by proposing the use of state-of-the-art configurations of the U-Net model, such as the Attention U-Net (Oktay et al., 2018) and Dynamic RU-NEXT (Jasmine and Marichamy, 2024) models, to identify AMPs from low-resolution micrographs. The Dynamic RU-NEXT model, introduced as an advanced framework for highly accurate tumor segmentation in CT scans, employs a refined architectural design to improve segmentation precision. This work also differs from previous studies as classification is directly integrated within the U-Net-based segmentation frameworks to streamline the workflow. The models are compared with the Mask R-CNN model alongside the Otsu thresholding method. The main contributions of this research include: (1) using micrographs of AMPs collected from outdoor environments as part of a pilot study using existing pollen monitoring devices, (2) creating the first labelled dataset of outdoor AMPs for image segmentation, to the best of the authors' knowledge; and (3) presenting the first use of deep learning for automated pre-screening of outdoor AMPs from low-resolution micrographs and comparing the results with traditional methods.

2. Materials and methods

2.1. Study locations

This study utilised images of AMPs collected outdoors using existing monitoring techniques. Data collection was part of a pilot study by Chapman et al. (2024) that used existing pollen monitoring equipment to monitor AMPs, making this one of the first studies to do so. Sampling was conducted using a Burkard pollen trap, which continuously collected air samples over two 7-day periods from the roof of the Hardy Building, University of Hull, U.K. (53° 46′ 16.87″ N; 0° 22′ 2.64″ W) (Google Maps, 2025). Sampling began at 10:00 AM on 10/03/2023 and 11:00 AM on 18/07/2023. Additionally, a second Burkard trap was deployed for 7 days at Nelson Mandela University's South Campus, Summerstrand, Gqeberha, S.A. (34° 0′ 4.66″ S; 25° 40′ 2.40″ E), commencing at 1:00 PM on 03/08/2023. Both traps were positioned on flat rooftops, with the sampling orifice at least 1 m above the roof surface and at least 2 m from the building edge (Fig. S1). Both locations were chosen as those already had a functional Burkard trap set up for sample collection.

2.2. Data preparation

The collected particles were imaged using a Nicolet iN10 μFTIR microscope which has a 15×0.7 numerical aperture high efficiency and condenser, and independent reflection and transmission illuminations. The photos were taken of samples on anodiscs so that it is clear what the background material in the photo is made of. Only images of particles whose spectra showed a $\geq 70\%$ match with those in the Omnic Picta and Omnic Polymer libraries were selected for this study. The sampled particles found in the images were predominantly polytetrafluoroethylene (PTFE), nylon, low-density polyethylene (LDPE), polystyrene (PS), poly(ethylene-propylene-diene) (EPDM), and nylon 6 (PA) (Table S1). More details on the collection, preparation, and
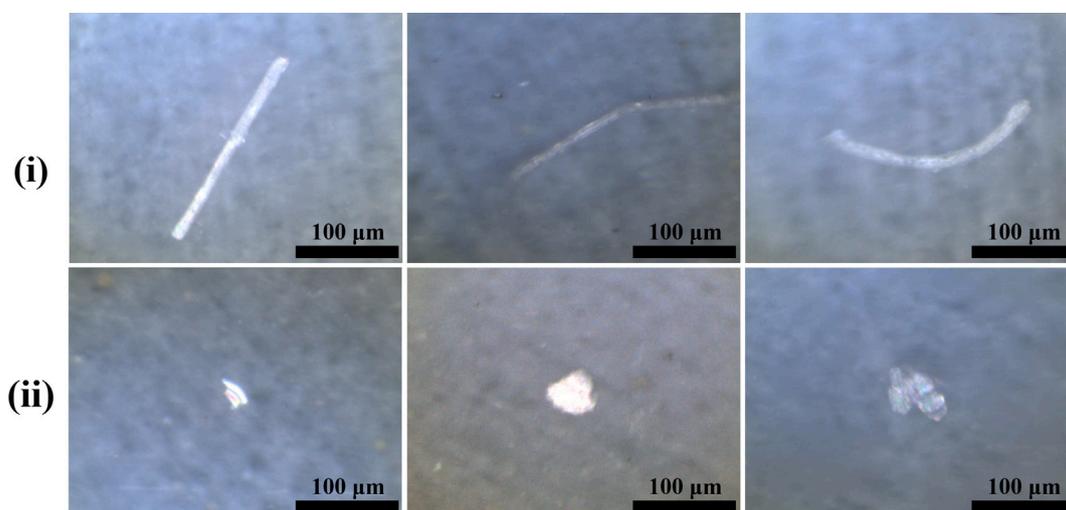
Fig. 1. Sample images of two shapes of AMPs: (i) fibre, and (ii) fragment.

chemical identification of the particles are provided in Chapman et al. (2024). The particles were categorised as either 'fragment' (irregularly shaped resulting from the breakdown of larger plastic debris (Free et al., 2014)), or as 'fibre' (with a length-to-width ratio > 3 (Vianello et al., 2019)). Fig. 1 presents sample images of the two AMP shapes utilised in this study. Fibres and fragments are the most common shapes of AMPs, varying based on the specific study examined (Table S2). Although prevalent in other studies, foam and spheres/microbeads were not observed in any samples.

The dataset reveals distinct patterns in the distribution of size, colour, and shape of the AMPs found in the micrographs. The size distribution of the particles exhibits high variability, with an average particle length of 85.17 ± 88.43 μm (range 18–750 μm), and an average particle width of 44.04 ± 33.60 μm (range 10–361 μm) (Fig. S2). The images also comprise 8 distinct colours, with 'clear' being the most dominant (approx. 69.85% of the total). Other colours, such as 'grey', 'white', and 'black', appear significantly less frequently (Fig. S3).

2.3. Data augmentation and splitting

Due to the limited collected samples, data augmentation techniques were applied to the images containing the sampled outdoor AMPs to address data imbalance and model overfitting. Vertical and horizontal flipping were applied to the samples, resulting in a balanced dataset containing 300 images per shape class. Each image was then manually annotated using an online tool (Makesense.ai, 2024) to create ground-truth masks for model training. The annotations were exported in COCO JSON format. A total of 10% of the dataset was used as testing data to allocate samples that were completely unknown to the models during training. A 5-fold cross-validation setup was then employed which involves splitting the training dataset into five equal batches. Each batch served as the validation set while the rest were allocated for model training. The process was repeated five times, and the overall performance was averaged across the models. The choices for selecting the testing set ratio and the number of cross-validations align with those of previous works (Lorenzo-Navarro et al., 2020; Shi et al., 2022; Huang et al., 2023). Data splitting was conducted with a fixed random seed to ensure reproducibility. Table 1 provides a distribution of the images before and after data augmentation.

2.4. Model development

2.4.1. U-Net model and its enhanced configurations

The U-Net architecture is a standard encoder–decoder framework for image segmentation, comprising a contracting path and an expansive path, as illustrated in Fig. 2(a). The contracting path serves as

Table 1
Composition of the dataset containing sampled outdoor AMPs.

		Total	Training	Validation	Testing
Collected images		272	196	49	27
Images with corresponding	Fibre	69	50	12	7
AMPs shapes	Fragment	203	146	37	20
Augmented images	Fibre	231	166	42	23
	Fragment	97	70	17	10
Total images		600	432	108	60

an encoder, extracting hierarchical features from input images through stacks of convolutional layers activated by Rectified Linear Unit (ReLU) functions, followed by max-pooling layers to reduce spatial dimensions. The expansive path acts as a decoder, reconstructing pixel-level segmentations via upsampling layers implemented with transposed convolutions and ReLU activation. Features extracted by the contracting path are concatenated with those in the expansive path using skip connections, preserving spatial information and enabling fast, precise segmentation (Ronneberger et al., 2015). A softmax function at the output assigns probability values to each pixel, determining whether it belongs to a specific class or the background.

Building upon the standard U-Net framework, the Attention U-Net and Dynamic RU-NEXT models introduce advanced modifications that further refine and enhance the segmentation process. The Attention U-Net also uses a contracting path to capture hierarchical features and an expansive path to reconstruct the segmentation map. However, it incorporates attention gates within the skip connections to highlight the most salient features. Instead of passing all encoder features indiscriminately to the decoder, these attention mechanisms selectively weight critical regions, suppressing irrelevant background information. The Dynamic RU-NEXT model builds upon these advancements with a residual U-shaped architecture and dynamic convolutional kernels that adapt to input data during runtime. In its decoder component, the model employs a dynamic U-Net structure with pixel shuffle upsampling, allowing for the generation of high-resolution segmentation maps from the extracted features (Jasmine and Marichamy, 2024). By doing so, the model can handle greater variability in input appearances, refining its feature extraction and segmentation strategies simultaneously. The Attention U-Net and Dynamic RU-NEXT models are thoroughly described in Oktay et al. (2018) and Jasmine and Marichamy (2024), respectively.

In this study, the architectures of the U-Net-based models were further extended to include a secondary output branch dedicated to classifying the entire image based on the segmented shapes. This was
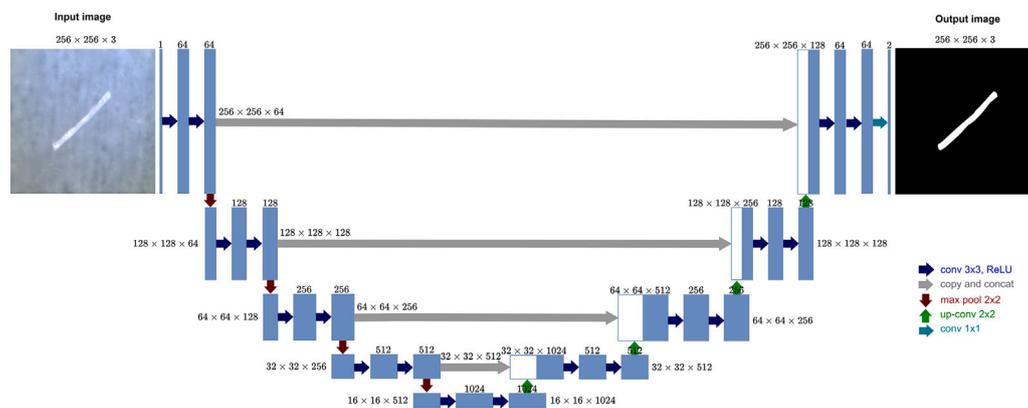
(a)
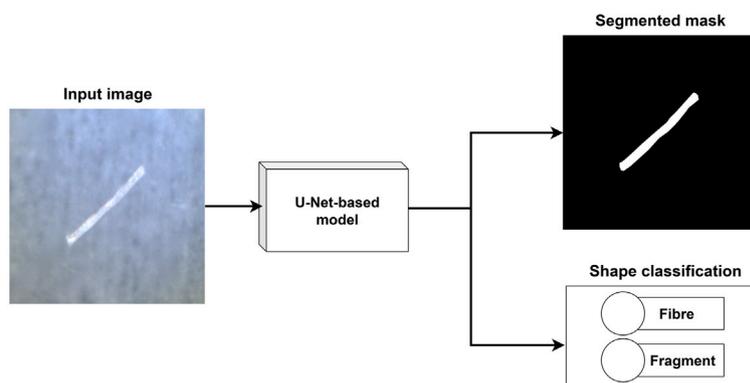
(b)

Fig. 2. (a) The U-Net model architecture (adapted from Ronneberger et al. (2015)); and (b) the proposed AMPs identification approach based on an enhanced U-Net architecture.

achieved by introducing a classification head that operates in parallel with the segmentation output, enabling simultaneous segmentation and classification (see Fig. 2(b)). This approach distinguishes this work from other related studies where classification was performed independently of segmentation. For instance, Lorenzo-Navarro et al. (2021) conducted classification separately from the segmentation of images containing MPs collected from marine environments, utilising U-Net and VGG16 models in a two-stage process. Similarly, Shi et al. (2022) adopted a parallel strategy by employing U-Net and MultiResUnet models for segmentation alongside a pre-trained VGG16 model for classification. By integrating classification directly within the segmentation framework, this study's approach streamlines the workflow, reduces computational overhead, and enhances the ability of the enhanced U-Net models to utilise shared feature representations for improved overall performance.

2.4.2. Mask R-CNN model

The Mask R-CNN model incorporates a ResNet-101 backbone which is a deep residual neural network that extracts detailed features from input images and produces rich feature maps (He et al., 2016). These feature maps are passed through region proposal networks (RPNs), which generate potential object regions by identifying regions of interest (ROIs) within bounding boxes, refined using anchor scaling and non-maximum suppression. ROI Align is applied to ensure the ROIs have uniform dimensions, enabling efficient processing. A three-branch prediction network handles classification, object localisation, and instance segmentation, with the classification and bounding box

refinement performed through fully connected layers and softmax functions. Instance segmentation is achieved using a fully convolutional network (FCN) that generates binary masks, accurately delineating detected objects at the pixel level. The Mask R-CNN model is thoroughly described in He et al. (2016).

2.5. Model training

The three U-Net-based models were trained using the Adam optimiser for 50 epochs, with a learning rate of 10^{-3}, momentum of 0.90, and a batch size of 8. Additionally, the models employed transfer learning by utilising encoders based on neural networks pre-trained on the ImageNet dataset, which comprises 14 million images spanning a wide array of object categories. Transfer learning is a widely adopted technique in deep learning as it reduces training time and is particularly effective when working with limited training data. In contrast, the Mask R-CNN model was trained using the stochastic gradient (SGD) algorithm for the same number of epochs, with a learning rate of 10^{-3}, momentum of 0.90, and a batch size of 2. Like the U-Net models, Mask R-CNN also benefited from transfer learning utilising weights from a model pre-trained on the COCO dataset, which includes over 300,000 images across 80 object categories (Lin et al., 2015). The hyperparameter choices are consistent with those in previous related works (Table S3).

Prior to model training, all models underwent grid optimisation to fine-tune their hyperparameter values which took around 8 h to complete per model. The training process was implemented using
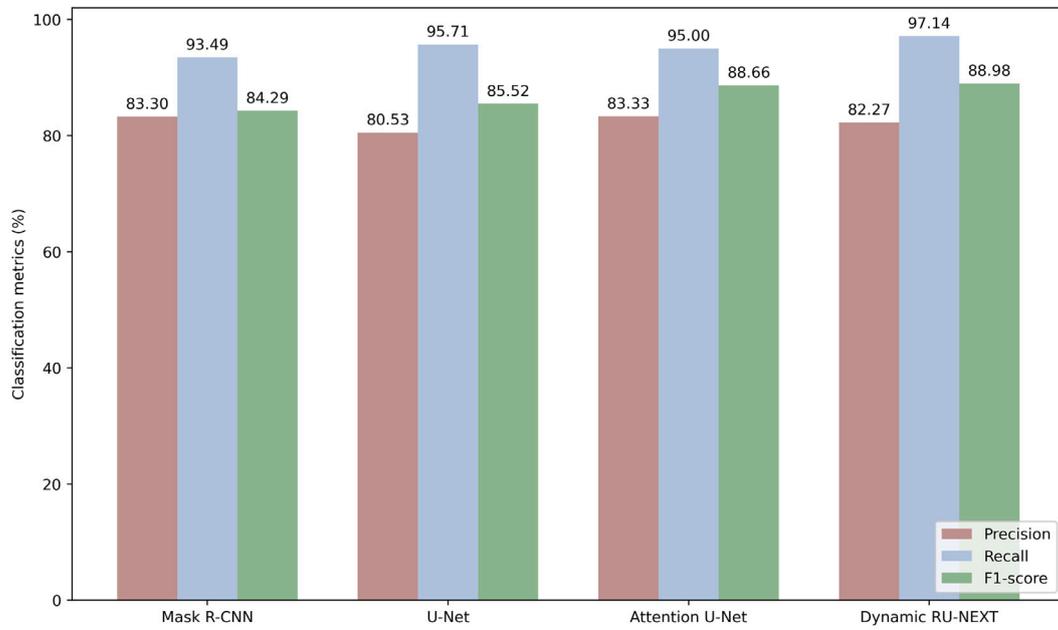
(a)
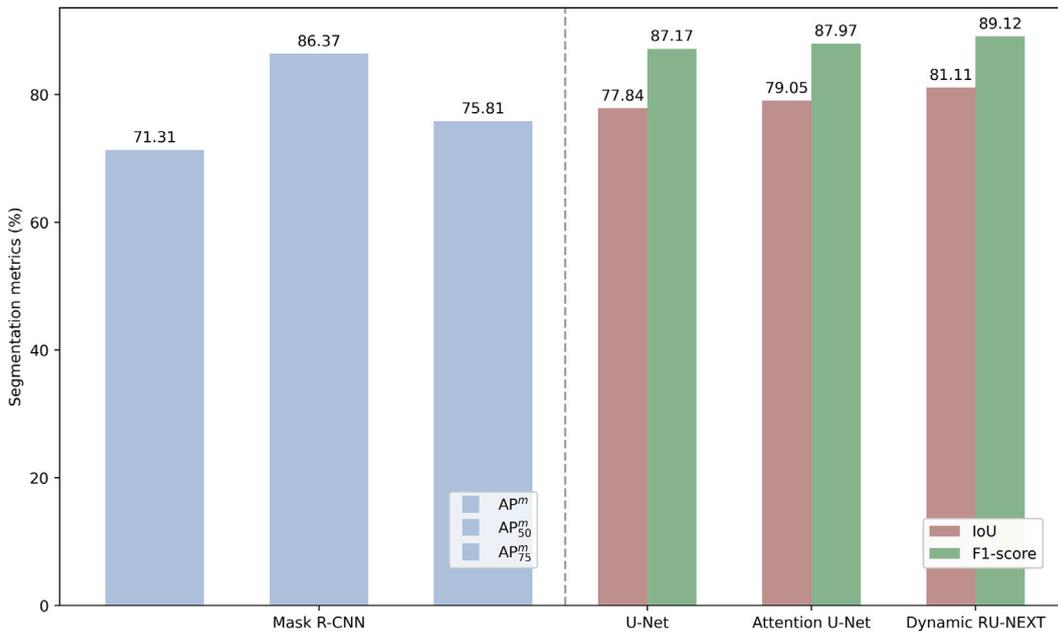
(b)

Fig. 3. Performance evaluation of the deep learning models: (a) classification, and (b) segmentation.

TensorFlow (version 2.9.3) in Python, supported by the Nvidia CUDA and cuDNN open-source toolkits, and managed via Anaconda (version 3.24.0). Model training was carried out on a personal computer running Windows 10, equipped with an AMD Ryzen 9 6900HX processor, 64 GB of DDR3 RAM, and an NVIDIA GeForce RTX 3080 Ti graphics card.

2.6. Model evaluation

2.6.1. Loss function

The loss function serves as a metric for evaluating the difference between predicted outputs and the ground truth. The selection of suitable loss functions is crucial during the training process to ensure

effective model calibration. For the enhanced U-Net-based models with classification, the loss function comprises two components:

$$L = L_{\text{class}} + L_{\text{seg}}. \tag{1}$$

The two terms in Eq. (1) are calculated as follows:

$$L_{\text{class}}\left(p_i, p_i^*\right) = -p_i^* \log p_i - \left(1 - p_i^*\right) \log \left(1 - p_i\right) \tag{2}$$

$$L_{\text{seg}} = = 1 - \frac{2 \times |\text{prediction result} \cap \text{ground truth}| + \lambda}{|\text{prediction result} \cup \text{ground truth}| + \lambda}, \tag{3}$$

where p_i and p_i^* represent the predicted and ground truth class probabilities, respectively, and λ is a term introduced to prevent the numerical issue of division by zero. L_{seg} is represented by the Dice loss

which was derived from the Dice coefficient to evaluate segmentation performance.

For the Mask R-CNN model, the loss function comprises three components since it simultaneously performs localisation, classification, and segmentation:

$$L = L_{\text{box}} + L_{\text{class}} + L_{\text{seg}}, \tag{4}$$

where L_{box}, L_{class} and L_{seg} represent the localisation error of the bounding box, classification loss, and segmentation loss of mask based on pixel accuracy, respectively. The three terms in Eq. (4) are calculated as follows:

$$L_{\text{box}}\left(t_i, t_i^*\right) = L_1 smooth \left(t_i - t_i^*\right) \tag{5}$$

$$L_1 smooth(x) = \begin{cases} 0.5x^2, & \text{if}|x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \tag{6}$$

$$L_{\text{mask}}\left(s_{ij}, s_{ij}^*\right) = -\frac{1}{m^2} \sum_{ij} \left[s_{ij}^* \log s_{ij}^k + \left(1 - s_{ij}^*\right) \log \left(1 - s_{ij}^k\right) \right] \tag{7}$$

where t_i is the predicted vector representing the bounding box's location and size, while t_i^* is the ground truth vector, m^2 refers to the mask resolution, typically 28×28 pixels, s_{ij} and s_{ij}^* denote the binary values (0 or 1) in the predicted and ground-truth masks, respectively, and k represents the kth class object in the dataset. In the segmentation task, the dataset is assumed to include two object categories ($k = 2$), fibre and fragment. These formulas ensure that the model effectively balances its performance across localisation, classification, and segmentation objectives.

2.6.2. Segmentation and classification metrics

To evaluate the performance of the deep learning models, several segmentation and classification metrics were utilised. The precision, recall, and F1-score metrics assess classification performance, and are defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{9}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{10}$$

where TP, FP, and FN denote true positives, false positives, and false negatives, respectively. In the context of fibre AMPs classification, TP is the number of pixels correctly classified as fibre, FP is the number of pixels incorrectly classified as fibre when they belong to the fragment class, and FN is the number of pixels belonging to fibres that were incorrectly classified as fragments. These metrics ensure a comprehensive evaluation of the models' ability to correctly identify fibre and fragment classes, capturing both its accuracy (precision) and completeness (recall).

The Intersection over Union (IoU) metric was employed to evaluate the accuracy of segmentation and localisation performance. IoU is defined as:

$$\text{IoU} = \frac{\text{area of overlap}}{\text{area of union}}, \tag{11}$$

where the area of overlap is the intersection between the predicted and the ground-truth masks or bounding boxes, and the area of union is the total area covered by both the predicted and ground-truth annotations.

2.6.3. Localisation metrics

For Mask R-CNN, the algorithm generates bounding boxes to indicate the locations of target objects for detection. The accuracy of these bounding boxes serves as a measure of localisation performance. However, before evaluating the proposed algorithm, it is essential to establish the criteria for correct localisation. In this study, the area of overlap, quantified by IoU between the ground truth and predicted bounding boxes, was used as the criterion for quantifying correct localisation. Following the approach adopted in related works, the average

precision of the bounding box (AP^{bb}) and its two variants, AP_{50}^{bb} and AP_{75}^{bb} (Han et al., 2023; Nie et al., 2020), were employed for evaluation. AP_{50}^{bb} and AP_{75}^{bb} represent the average precision when the IoU is 50% and 75%, respectively, while AP^{bb} denotes the mean value of average precision scores calculated for IoU thresholds ranging from 50% to 95% in increments of 5%. This is consistent with the findings of Han et al. (2023). The values for all metrics range between 0 and 1, with 0 indicating no agreement between the models' predictions and the ground truth and 1 representing perfect agreement. Metrics values were reported with four significant digits to account for differences smaller than 0.1%.

3. Results

3.1. Localisation performance

The Mask R-CNN model achieved an average values of AP_{50}^{bb}, AP_{75}^{bb}, and AP^{bb}, ranging from 73.92% to 91.09% on the test images via 5-fold cross-validation (Table S4). These results demonstrate the model's strong localisation performance, particularly under the less stringent IoU threshold of 50% (AP_{50}^{bb}). However, the performance decreases as the IoU threshold increases, with AP_{75}^{bb} and AP^{bb} capturing stricter localisation requirements. The drop in AP_{75}^{bb} to 78.92% and the even lower overall average of 73.32% for AP^{bb} is expected, as higher IoU thresholds demand more precise alignment between the predicted and ground truth bounding boxes.

3.2. Classification performance

As illustrated in Fig. 3(a), the U-Net model attained average precision, recall, and F1-scores of 80.53%, 95.71%, and 85.52%, respectively (Table S3). The integration of attention mechanisms into Attention U-Net resulted in a slight improvement in precision and F1-score values by approx. 3%, while sustaining a high recall. The Dynamic RU-NEXT model further enhanced performance by achieving the highest recall of 97.14% and an F1-score of 88.98%, alongside a precision of 82.27%. This superior performance of Dynamic RU-NEXT can be attributed to its advanced architecture, which effectively combines residual connections and attention mechanisms to enhance feature extraction. Consequently, Dynamic RU-NEXT exhibits an enhanced ability to accurately identify positive instances without compromising overall performance.

The Mask R-CNN model achieved average precision, recall, and F1-score of 83.30%, 93.49%, and 84.29%, respectively (Table S5). These results exhibit the model's robust capability for accurate object detection, with a high recall highlighting its effectiveness in identifying the majority of target objects present in the images. The Mask R-CNN model also achieved more accurate classification results than those reported in Huang et al. (2023), where substantial differences in model performances across shape classes were observed. This can be attributed to the highly variable dataset the study utilised, which included over 4000 images from their samples and over 200 articles retrieved from Google Scholar. Furthermore, the classification performance in this study is comparable to that of Han et al. (2023) for images with a clean background, and outperforms their results on images with natural backgrounds. Other prior studies, such as Lorenzo-Navarro et al. (2021) and Shi et al. (2022), employed the VGG16 model with transfer learning, outperforming the classification results of this work (average precision and recall values of approx. 98%). However, a direct comparison cannot be made as these studies conducted segmentation and classification in two independent phases.

Fig. 4(a) presents representative prediction outputs generated by the Mask R-CNN model, where (i) and (ii) display sample predictions on micrographs containing fibre- and fragment-shaped AMPs, respectively. From visual inspection, the Mask R-CNN model achieved good performance in detecting AMPs with high classification scores across different scenarios. In contrast, Fig. 4(b) highlights a selection of particle shapes

that the Dynamic RU-NEXT model misclassified. Certain fibre-shaped AMPs were classified as fragments due to their irregular edges, while some fragment-shaped AMPs with comparatively smoother edges were mistaken for fibres.

3.3. Segmentation performance

The segmentation results of the U-Net-based models on the test images are shown in Fig. 3(b). The U-Net, Attention U-Net, and Dynamic RU-NEXT models achieved IoU scores of 77.84%, 79.05%, and 81.11%, respectively (Table S6). These results suggest that incorporating attention mechanisms and dynamic convolutions, as well as lightweight transformer-inspired modules, into the U-Net architecture significantly improved its segmentation performance.

The results of this study are lower than those reported in Han et al. (2023), where an average IoU score of 87.30% was achieved after U-Net models were trained on high-resolution images of MPs against a white background. However, their model performed poorly when trained on images with sand, natural soil, and water backgrounds, yielding an IoU score of 10.80%. Similarly, Shi et al. (2022) obtained slightly higher cross-validated IoU scores (approx. 89%) than this study, likely due to their use of SEM imaging, which produced uniform black or grey backgrounds. Lorenzo-Navarro et al. (2021) also reported a higher U-Net score (80%) by capturing high-resolution images of MPs on DIN-A4 paper, ensuring a clean white background. However, the Dynamic U-Next model in this study outperformed their results by approx. 1.39%. In contrast, this study reports superior segmentation performance compared to the U-Net models trained in Huang et al. (2023), which achieved lower F1-scores (approx. 41%). This suboptimal performance may be due to the highly diverse image dataset used for training, which exhibited a significant class imbalance.

The metrics AP_{50}^m, AP_{75}^m, and AP^m were also utilised to evaluate the segmentation results of the Mask R-CNN model on the test images. Table S4 demonstrates the model's capability to perform segmentation tasks effectively, with AP_{50}^m value of 86.37% indicating a high success rate when a less stringent IoU threshold of 50% is applied. However, the performance decreases when stricter criteria are applied. For instance, the AP_{75}^m value of 75.81% reflects the model's segmentation accuracy under a more demanding IoU threshold of 75%, while the overall AP^m, averaging across multiple IoU thresholds from 50% to 95%, drops further to 65.11%. This decline underscores the challenge of achieving precise pixel-level segmentation, especially in images with objects in low-contrast and uneven backgrounds. The results of this study are consistent with those of Han et al. (2023), although their Mask R-CNN model achieved a significantly higher $AP^{(m)}$ score of 82.60% when trained on a clear background and a lower score of 59.50% when trained on natural backgrounds.

3.4. Comparison of the deep learning models vs. Otsu thresholding method

The models were also evaluated against a traditional thresholding-based approach known as the Otsu method. This method automatically determines a global threshold value based on the intensity distribution of the image. Before applying Otsu's thresholding, the images were converted to grayscale and subjected to Gaussian blurring using a kernel size of 11 × 11 to reduce noise and enhance segmentation accuracy. The binarised segmentation outputs were then inverted to account for the higher intensity of the background relative to the target objects.

Fig. 5(a) presents sample predictions generated by the Mask R-CNN, Dynamic RU-NEXT models, and Otsu thresholding method. The visual results align with the calculated IoU values and AP_{75}^m scores exceeding 75%. The models effectively identified fibre and fragment particles even in low-contrast and uneven image backgrounds. However, both models struggled to capture fine details at the edges of fragment particles, as illustrated in Fig. 5(a)(ii).
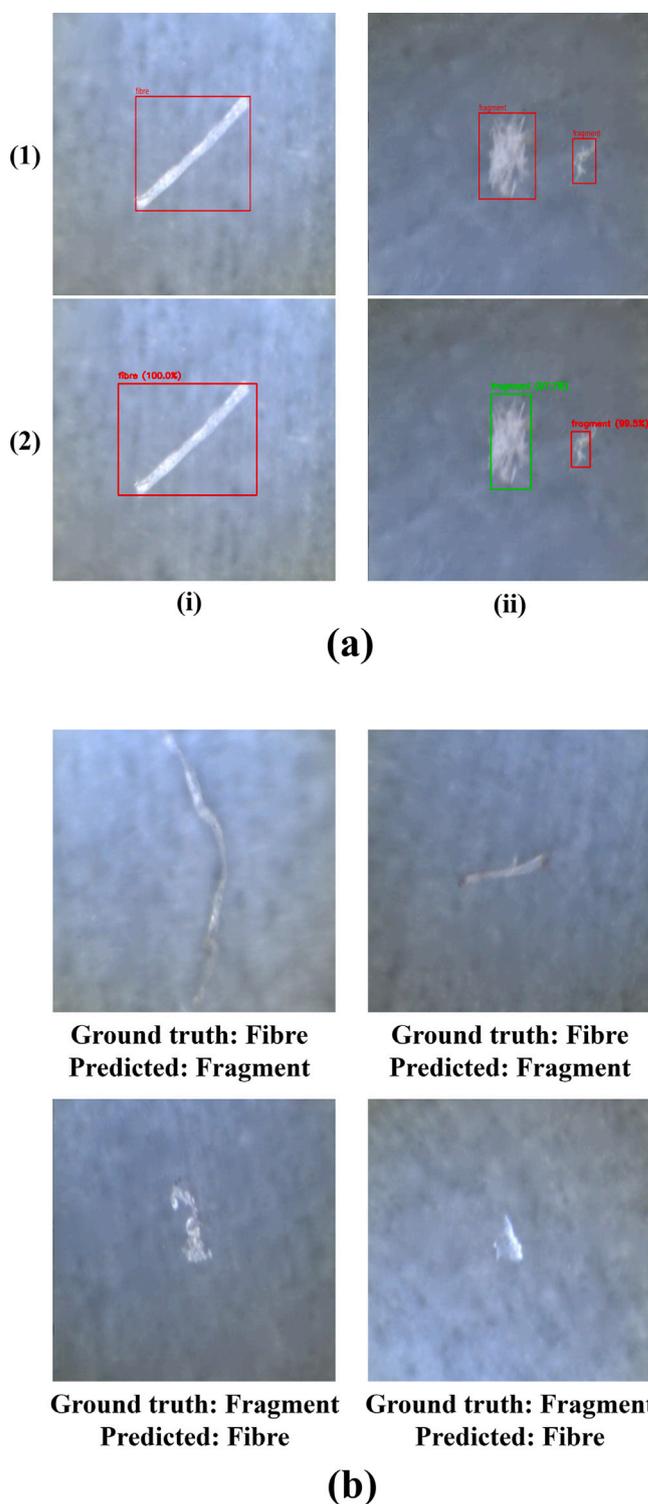
Fig. 4. (a) Sample localisation and classification results of the Mask R-CNN model on micrographs containing: (i) fibre; (ii) two fragments, where (1) and (2) indicate the ground-truth and predicted bounding box and shape class, respectively, and (b) Examples of misclassified AMPs by the Dynamic RU-NEXT model.

The Otsu method performs well on images exhibiting strong contrast between the target object and the background, but it is unsuitable for images where the object is nearly transparent or the background contains numerous artefacts. This limitation leads to a low IoU score of 32.14% on the test images, highlighting its inadequacy for complex segmentation tasks. Shi et al. (2022) also reported similar challenges
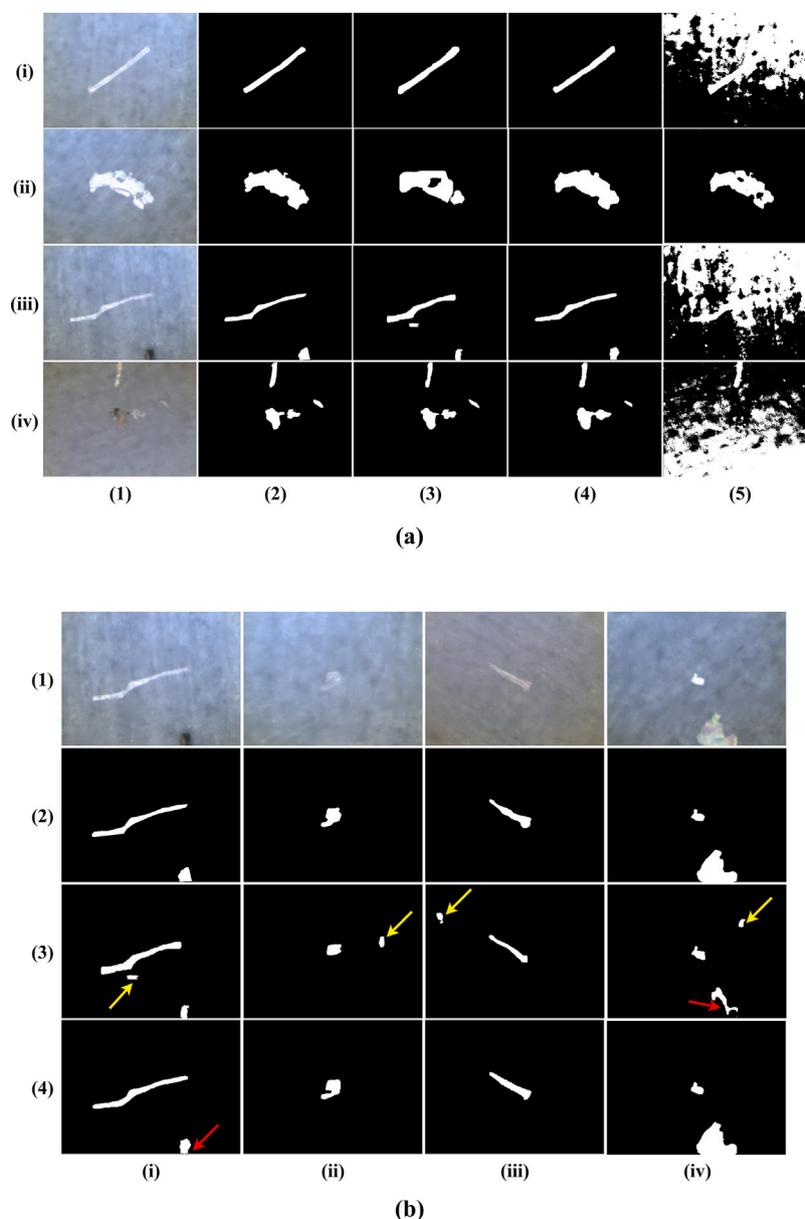
Fig. 5. (a) Sample segmentation results: micrographs with (i) fibre, (ii) fragment, (iii) a fibre and a fragment, and (iv) three fragments and a fibre; and (b) Sample results illustrating false and missed detections: micrographs with (i, iii) fibre and (ii, iv) fragment. (1) original images, (2) ground-truth masks, images produced by (3) Mask R-CNN, (4) Dynamic RU-NEXT, and (5) Otsu thresholding.

when applying the Otsu method to identify MPs on SEM images. Although fine-tuning can improve the Otsu method's performance for specific images, its sensitivity and reliance on contrast make it less effective for large and diverse datasets.

Fig. 5(b) presents sample predictions from the Dynamic RU-NEXT and Mask R-CNN models, highlighting instances of both false positives and missed AMPs. The yellow and red arrows indicate false positives and false negatives, respectively. The false detections could result from the transparency of the target particles, which makes them difficult to distinguish from the background, as well as the subtle contrast between the particles and their surroundings. Additionally, the presence of background noise, such as dirt and artefacts in the sample images, introduces further challenges for accurate identification.

3.5. Computational efficiency of the deep learning models

To assess the computational efficiency of the models, the total training times were summarised in Table 2. These values reflect the

average training duration per fold during cross-validation. The U-Net-based models have significantly fewer trainable parameters compared to the Mask R-CNN model, with parameter counts ranging from 11 to 18 million. This reduced parameter count results in shorter training durations and faster per-image processing times compared to the Mask R-CNN model. For instance, the U-Net model demonstrates superior efficiency with significantly fewer parameters, a shorter training time, and the fastest image processing speed. In contrast, the Mask R-CNN model has significantly more trainable parameters (approx. 64.6 million) despite utilising transfer learning. This can be attributed to the complexity of tasks it performs, such as localisation, classification, and segmentation. Consequently, it requires an average of 4715.14 s to complete its training process. This is also reflected in the model's average image processing time which is 0.36 s per image. Additionally, Table 2 presents the average of the number of epochs it took for each model to achieve the best performance. Although training was conducted for a maximum of 50 epochs, the epoch associated with the model weights exhibiting the best performance was saved during the

Table 2

Computational efficiency of the deep learning models.

Model	Parameter count	Maximum epochs	Best epochs	Training time (sec.)	Processing time per image (sec./image)
Mask R-CNN	64,674,852	50	45.4	4715.14	0.3593
U-Net	11,305,348	50	39	625.48	0.0465
Attention U-Net	11,654,952	50	41.6	760.07	0.0484
Dynamic RU-NEXT	18,630,632	50	43.6	1560.91	0.0489

training process. The results indicate that the models achieved optimal performance even before reaching the maximum number of epochs.

A direct comparison of segmentation and classification times across studies reveals that the proposed approach of this work is among the most computationally efficient. For segmentation, the Mask R-CNN model reported by Huang et al. (2023) achieved an average processing time of 0.20 s per micrograph, whereas the U-Net-based models in this work further reduced this to 0.05 s. Similarly, Shi et al. (2022) reported that their trained U-Net models required only a few seconds to segment a micrograph, aligning with the findings of this work. However, in prior studies, segmentation and classification were conducted separately, requiring additional computational steps for shape identification. For classification, Lorenzo-Navarro et al. (2021) employed a VGG16-based model, which took an average of 7 s to classify a micrograph which is significantly longer than this work's integrated approach.

4. Discussion

This study employed a range of deep learning models, including U-Net, Attention U-Net, Dynamic RU-NEXT, and Mask R-CNN, to segment and classify AMPs in micrographs. These models were trained on images containing AMPs with varying shapes, sizes, and colours to evaluate their adaptability to diverse particle morphologies.

The Mask R-CNN model demonstrated strong segmentation performance, particularly under less stringent conditions, with a notable AP_{50}^{m} of 86.37%. However, its low AP^{m} score of 65.11% highlights the inherent difficulty of achieving precise pixel-level segmentation in images with uneven backgrounds and those containing small-sized AMPs. All three U-Net configurations achieved IoU scores which met the widely recognised segmentation benchmark of IoU > 70% (Lin et al., 2015). The Dynamic RU-NEXT model provided the most accurate results which can be attributed to its incorporation of dynamic convolutions and attention mechanisms. These enhancements underscore its ability to focus on relevant features and maintain high segmentation accuracy across diverse scenarios.

Since the images in this study predominantly contained a single AMPs shape category, the integrated classification feature within the enhanced U-Net-based models provided a more efficient and accurate solution than the bounding-box-based Mask R-CNN. In environments where multiple particle shapes coexist in a single image, a more comprehensive model such as Mask R-CNN might be preferred.

Crucially, by integrating classification directly within the U-Net-based models, Attention U-Net and Dynamic RU-NEXT not only excelled in segmentation but also achieved high classification results, with F1-scores over 85%. This integration streamlines the workflow, eliminating the need for separate classification stages and enhancing computational efficiency. Consequently, these enhanced U-Net models offer a more efficient and effective approach to AMPs identification and classification compared to traditional methods.

Otsu thresholding method served as a baseline for comparison. While effective in high-contrast scenarios, the Otsu method struggled with the variability present in this study's images (IoU < 70%), such as low-contrast or transparent particles and uneven background colours. In contrast, the deep learning models proved robust, requiring only ground-truth masks for effective training and demonstrating consistent segmentation performance across diverse imaging conditions.

This work also evaluated the Segment Anything Model (SAM), which has rapidly emerged as a leading framework for zero-shot learning on new datasets (Kirillov et al., 2023). Zero-shot learning refers to the ability of a model to undertake new tasks without having been trained on labelled examples specifically for those tasks. Unlike traditional segmentation approaches that require large amounts of domain-specific data for fine-tuning, SAM can directly segment unseen images through prompt-based instructions. SAM was applied to the test images of this study without any prior training and still achieved impressive IoU and F1-score values (73.80% and 82.11%, respectively). This performance underscores SAM's adaptability and highlights its potential to drastically reduce the need for extensive labelled datasets. Sample segmentation results of SAM can be found in Fig. S4. The continued exploration of SAM's flexible segmentation capabilities could pave the way for innovative applications, such as combining SAM with other advanced deep learning frameworks or refining prompt-based interactions to tackle increasingly complex segmentation scenarios.

The models performed impressively on low-resolution images (256 × 256 pixels), contrasting with previous studies requiring higher resolutions for acceptable performance (Han et al., 2023; Lorenzo-Navarro et al., 2021; Shi et al., 2022). This suggests that deep learning methods can be optimised for use on less powerful computing platforms, broadening their applicability for resource-constrained settings.

The proposed approach offers significant potential as a pre-screening tool for detecting MPs in diverse environmental samples. The models in this study demonstrated adaptability to varying background colours, particle shapes, and particle-to-background size ratios, making them suitable for application across different imaging conditions. However, practical considerations must be addressed to ensure effective deployment in real-world scenarios. Pre-processing steps, such as the removal of unwanted particles (e.g., dirt and sand), and manual annotation for generating ground-truth masks remain necessary. For instance, annotation can be time-intensive, requiring approx. 5 min per micrograph. Additionally, model training requires computers equipped with GPUs to reduce training time, which can range from 13 to 80 min depending on hyperparameter configurations. Despite these requirements, fully-trained models can process an image in less than a second. This enables their deployment across local workstations or cloud platforms, even those without GPU support.

Integrating the proposed deep learning approach with complementary chemical analysis techniques, such as FTIR or Raman spectroscopy, presents a holistic solution for AMPs identification. While spectroscopic methods are highly accurate for confirming chemical composition, they are labour-intensive and unsuitable for high-throughput screening. By using deep learning models to pre-screen micrographs, researchers can prioritise particles likely to be MPs for further chemical characterisation, significantly reducing the time required for large-scale analyses. The findings of this study further underscore the potential of machine learning to systematically monitor emerging environmental pollutants, especially outdoor MPs, as recommended by other recent studies (Cui et al., 2023; Zhen et al., 2023; Zhu et al., 2023; Withana et al., 2024; Gaur et al., 2024; Zhao et al., 2024; Fazil et al., 2024).

Despite the significant results of this work, several limitations and challenges were encountered that future efforts could address:

1. Chemical analysis of MPs remains crucial for understanding their environmental and health effects. Features such as aerodynamic

diameter distribution, particle size (e.g., Feret diameter), and degree of degradation are also vital for precise MPs characterisation. Future research could explore incorporating these parameters into existing deep learning frameworks to enhance their comprehensiveness.

2. This study focused solely on fibres and fragments, as these are the only shapes of AMPs that were collected during sampling. However, other shapes such as films, spheres, and granules are also significant pollutants. Incorporating these additional shapes into future analyses would broaden the applicability of the proposed deep learning models.

3. The magnification used to capture MPs images likely played a significant role in the performance of the models in this study. The particles were imaged at a magnification of 100 μm, producing images with sufficient depth and detail for effective feature extraction. However, variations in imaging equipment or conditions may require different magnification ranges, potentially impacting model performance. Future studies should explore the effects of magnification on image quality and model outcomes, while also adapting particle pre-screening methods and deep learning architectures to account for these variations.

4. Smaller AMPs may present greater challenges in segmentation and classification due to their lower pixel resolution, increased edge ambiguity, and greater susceptibility to background noise. Future studies could further investigate the impact of particle size on model performance to refine classification accuracy and robustness.

5. Future studies might encounter challenges with increased irregular-edged particles and background noise, complicating image segmentation. This can lead to masks with extraneous background pixels and surrounding segmented particles. To address this, morphological operations like erosion and dilation can be employed to refine segmentation masks.

5. Conclusions

This study highlighted the capability of deep learning models to automate the pre-screening of AMPs in micrographs with high accuracy. The Mask R-CNN model offers comprehensive functionality for detailed instance-level analysis, while U-Net variants, including Attention U-Net and Dynamic RU-NEXT, provide efficient segmentation results enhanced with integrated classification capabilities. The creation of an open-source dataset, comprising images of AMPs with manually annotated masks, further contributes valuable resources to the field. By developing robust, pre-trained deep learning models capable of rapid deployment on local computers, the feasibility of real-time and large-scale monitoring of AMPs is significantly enhanced.

Integrating deep learning with established chemical characterisation techniques offers a comprehensive approach to AMPs identification. This further facilitates the development of a unified global method for characterising MPs and advancing our understanding of their environmental and health impacts.

CRediT authorship contribution statement

Sheen Mclean Cabaneros: Writing – original draft, Software, Methodology, Conceptualization. **Emma Chapman:** Writing – review & editing, Resources, Data curation. **Mark Hansen:** Writing – review & editing, Methodology. **Ben Williams:** Writing – review & editing. **Jeanette Rotchell:** Writing – review & editing, Supervision, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.envpol.2025.125993.

Data availability

Data will be made available on request.

References

Adhikari, S., Kelkar, V., Kumar, R., Halden, R.U., 2022. Methods and challenges in the detection of microplastics and nanoplastics: a mini-review. Polym. Int. 71 (5), 543–551. http://dx.doi.org/10.1002/pi.6348.

Bovik, A.C., 2009. Basic binary image processing. In: Bovik, A. (Ed.), The Essential Guide to Image Processing. Academic Press, Boston, pp. 69–96. http://dx.doi.org/10.1016/B978-0-12-374457-9.00004-4, (Chapter 4).

Chapman, E., Liddle, C.R., Williams, B., Hilmer, E., Quick, L.J., Garcia, A.G., Suárez, D.C., White, D., Bunting, M.J., Walker, P., Cabaneros, S.M.S., Kinnersley, R., Hansen, M.F., Atherall, C.A., Rotchell, J.M., 2024. Airborne microplastic monitoring: Developing a simplified outdoor sampling approach using pollen monitoring equipment. J. Hazard. Mater. 480, 136129. http://dx.doi.org/10.1016/j.jhazmat.2024.136129.

Chen, G., Feng, Q., Wang, J., 2020. Mini-review of microplastics in the atmosphere and their risks to humans. Sci. Total Environ. 703, 135504. http://dx.doi.org/10.1016/j.scitotenv.2019.135504.

Cui, S., Gao, Y., Huang, Y., Shen, L., Zhao, Q., Pan, Y., Zhuang, S., 2023. Advances and applications of machine learning and deep learning in environmental ecology and health. Environ. Pollut. 335, 122358. http://dx.doi.org/10.1016/j.envpol.2023.122358.

Danopoulos, E., Twiddy, M., West, R., Rotchell, J.M., 2022. A rapid review and meta-regression analyses of the toxicological impacts of microplastic exposure in human cells. J. Hazard. Mater. 427, 127861. http://dx.doi.org/10.1016/j.jhazmat.2021.127861.

Dewika, M., Markandan, K., Irfan, N.A., Mohd Abdah, M.A.A., Ruwaida, J., Sara, Y., Khalid, M., 2023. Review of microplastics in the indoor environment: Distribution, human exposure and potential health impacts. Chemosphere 324, 138270. http://dx.doi.org/10.1016/j.chemosphere.2023.138270.

Fazil, A.Z., Gomes, P.I., Sandamal, R.K., 2024. Applicability of machine learning techniques to analyze microplastic transportation in open channels with different hydro-environmental factors. Environ. Pollut. 357, 124389. http://dx.doi.org/10.1016/j.envpol.2024.124389.

Free, C.M., Jensen, O.P., Mason, S.A., Eriksen, M., Williamson, N.J., Boldgiv, B., 2014. High-levels of microplastic pollution in a large, remote, mountain lake. Marine Poll. Bull. 85 (1), 156–163. http://dx.doi.org/10.1016/j.marpolbul.2014.06.001.

Gauci, A., Deidun, A., Montebello, J., Abela, J., Galgani, F., 2019. Automating the characterisation of beach microplastics through the application of image analyses. Ocean & Coastal Management 182, 104950. http://dx.doi.org/10.1016/j.ocecoaman.2019.104950.

Gaur, V.K., Gautam, K., Vishvakarma, R., Sharma, P., Pandey, U., Srivastava, J.K., Varjani, S., Chang, J.S., Ngo, H.H., Wong, J.W., 2024. Integrating advanced techniques and machine learning for landfill leachate treatment: Addressing limitations and environmental concerns. Environ. Pollut. 354, 124134. http://dx.doi.org/10.1016/j.envpol.2024.124134.

Gonzales-Barron, U., Butler, F., 2006. A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumb features by digital image analysis. J. Food Eng. 74 (2), 268–278. http://dx.doi.org/10.1016/j.jfoodeng.2005.03.007.

Google Maps, 2025. University of Hull campus and Nelson Mandela University South campus satellite images. Google Maps [online]. (Accessed 7 February 2025).

Hale, R.C., Seeley, M.E., La Guardia, M.J., Mai, L., Zeng, E.Y., 2020. A global perspective on microplastics. J. Geophys. Res.: Ocean. 125 (1), http://dx.doi.org/10.1029/2018JC014719, e2018JC014719 2018JC014719.

Han, X.L., Jiang, N.J., Hata, T., Choi, J., Du, Y.J., Wang, Y.J., 2023. Deep learning based approach for automated characterization of large marine microplastic particles. Mar. Environ. Res. 183, 105829. http://dx.doi.org/10.1016/j.marenvres.2022.105829.

Hartmann, N.B., Hüffer, T., Thompson, R.C., Hassellöv, M., Verschoor, A., Daugaard, A.E., Rist, S., Karlsson, T., Brennholt, N., Cole, M., Herrling, M.P., Hess, M.C., Ivleva, N.P., Lusher, A.L., Wagner, M., 2019. Are we speaking the same language? Recommendations for a definition and categorization framework for plastic debris. Environ. Sci. Technol. 53 (3), 1039–1047. http://dx.doi.org/10.1021/acs.est.8b05297, pMID: 30608663.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.

Horvatits, T., Tamminga, M., Liu, B., Sebode, M., Carambia, A., Fischer, L., Püschel, K., Huber, S., Fischer, E.K., 2022. Microplastics detected in cirrhotic liver tissue. EBioMedicine 82, 104147. http://dx.doi.org/10.1016/j.ebiom.2022.104147.

Huang, H., Cai, H., Qureshi, J.U., Mehdi, S.R., Song, H., Liu, C., Di, Y., Shi, H., Yao, W., Sun, Z., 2023. Proceeding the categorization of microplastics through deep learning-based image segmentation. Sci. Total Environ. 896, 165308. http://dx.doi.org/10.1016/j.scitotenv.2023.165308.

Ibrahim, Y.S., Tuan Anuar, S., Azmi, A.A., Wan Mohd Khalik, W.M.A., Lehata, S., Hamzah, S.R., Ismail, D., Ma, Z.F., Dzulkarnaen, A., Zakaria, Z., Mustaffa, N., Tuan Sharif, S.E., Lee, Y.Y., 2021. Detection of microplastics in human colectomy specimens. JGH Open 5 (1), 116–121. http://dx.doi.org/10.1002/jgh3.12457.

Jahandari, A., 2023. Microplastics in the urban atmosphere: Sources, occurrences, distribution, and potential health implications. J. Hazard. Mater. Adv. 12, 100346. http://dx.doi.org/10.1016/j.hazadv.2023.100346.

Jasmine, S., Marichamy, P., 2024. Dynamic RU-Next: Advancing liver and tumor segmentation with enhanced deep learning architecture. J. Radiat. Res. Appl. Sci. 17 (4), 101182. http://dx.doi.org/10.1016/j.jrras.2024.101182.

Jenner, L.C., Rotchell, J.M., Bennett, R.T., Cowen, M., Tentzeris, V., Sadofsky, L.R., 2022. Detection of microplastics in human lung tissue using μFTIR spectroscopy. Sci. Total Environ. 831, 154907. http://dx.doi.org/10.1016/j.scitotenv.2022.154907.

Kim, Y.N., Yoon, J.H., Kim, K.H.J., 2020. Microplastic contamination in soil environment – a review. Soil Sci. Annu. 71 (4), 300–308. http://dx.doi.org/10.37501/soilsa/131646.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R., 2023. Segment anything. arXiv preprint arXiv:2304.02643, abs/2304.02643, (Accessed 06 June 2024).

Kroon, F., Motti, C., Talbot, S., Sobral, P., Puotinen, M., 2018. A workflow for improving estimates of microplastic contamination in marine waters: A case study from North-Western Australia. Environ. Pollut. 238, 26–38. http://dx.doi.org/10.1016/j.envpol.2018.03.010.

Leslie, H.A., van Velzen, M.J., Brandsma, S.H., Vethaak, A.D., Garcia-Vallejo, J.J., Lamoree, M.H., 2022. Discovery and quantification of plastic particle pollution in human blood. Environ. Int. 163, 107199. http://dx.doi.org/10.1016/j.envint.2022.107199.

Lin, J.Y., tao Liu, H., Zhang, J., 2022. Recent advances in the application of machine learning methods to improve identification of the microplastics in environment. Chemosphere 307, 136092. http://dx.doi.org/10.1016/j.chemosphere.2022.136092.

Lin, T.Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C.L., Dollár, P., 2015. Microsoft COCO: Common objects in context. arXiv:1405.0312.

Lorenzo-Navarro, J., Castrillón-Santana, M., Sánchez-Nielsen, E., Zarco, B., Herrera, A., Martínez, I., Gómez, M., 2021. Deep learning approach for automatic microplastics counting and classification. Sci. Total Environ. 765, 142728. http://dx.doi.org/10.1016/j.scitotenv.2020.142728.

Lorenzo-Navarro, J., Castrillón-Santana, M., Santesarti, E., De Marsico, M., Martínez, I., Raymond, E., Gómez, M., Herrera, A., 2020. SMACC: A system for microplastics automatic counting and classification. IEEE Access 8, 25249–25261. http://dx.doi.org/10.1109/ACCESS.2020.2970498.

Makesense.ai, 2024. MakeSense: Free online AI-powered image annotation tool. Website, URL: https://www.makesense.ai/. (Accessed 01 June 2024).

Mukhanov, V.S., Litvinyuk, D.A., Sakhon, E.G., Bagaev, A.V., Veerasingam, S., Venkatachalapathy, R., 2019. A new method for analyzing microplastic particle size distribution in marine environmental samples. Ecol. Montenegrina 23, 77–86. http://dx.doi.org/10.37828/em.2019.23.10.

Netema, B.N., Chakraborty, T.K., Nice, M.S., Islam, K.R., Debnath, P.C., Chowdhury, P., Rahman, M.S., Halder, M., Zaman, S., Ghosh, G.C., Islam, M.S., 2024. Appraisal of microplastic pollution and its related risks for urban indoor environment in Bangladesh using machine learning and diverse risk evolution indices. Environ. Pollut. 360, 124631. http://dx.doi.org/10.1016/j.envpol.2024.124631.

Nie, X., Duan, M., Ding, H., Hu, B., Wong, E.K., 2020. Attention mask R-CNN for ship detection and segmentation from remote sensing images. IEEE Access 8, 9325–9334. http://dx.doi.org/10.1109/ACCESS.2020.2964540.

Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M.C.H., Heinrich, M.P., Misawa, K., Mori, K., McDonagh, S.G., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D., 2018. Attention U-Net: Learning where to look for the pancreas. CoRR abs/1804.03999, arXiv:1804.03999.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 9 (1), 62–66. http://dx.doi.org/10.1109/TSMC.1979.4310076.

Pauly, J.L., Stegmeier, S.J., Allaart, H.A., Cheney, R.T., Zhang, P.J., Mayer, A.G., Streck, R.J., 1998. Inhaled cellulosic and plastic fibers found in human lung tissue. Cancer Epidemiol. Biomarkers Prevent. 7 (5), 419–428, pMID: 9610792.

Primpke, S., Christiansen, S.H., Cowger, W., Frond, H.D., Deshpande, A., Fischer, M., Holland, E.B., Meyns, M., O'Donnell, B.A., Ossmann, B.E., Pittroff, M., Sarau, G., Scholz-Böttcher, B.M., Wiggin, K.J., 2020. Critical assessment of analytical methods for the harmonized and cost-efficient analysis of microplastics. Appl. Spectrosc. 74 (9), 1012–1047. http://dx.doi.org/10.1177/0003702820921465, pMID: 32249594.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. CoRR, abs/1505.04597, arXiv:1505.04597.

Sauvola, J., Pietikäinen, M., 2000. Adaptive document image binarization. Pattern Recognit. 33 (2), 225–236. http://dx.doi.org/10.1016/S0031-3203(99)00055-2.

Shao, L., Li, Y., Jones, T., Santosh, M., Liu, P., Zhang, M., Xu, L., Li, W., Lu, J., Yang, C.X., Zhang, D., Feng, X., BéruBé, K., 2022. Airborne microplastics: A review of current perspectives and environmental implications. J. Clean. Prod. 347, 131048. http://dx.doi.org/10.1016/j.jclepro.2022.131048.

Shi, B., Patel, M., Yu, D., Yan, J., Li, Z., Petriw, D., Pruyn, T., Smyth, K., Passeport, E., Miller, R.D., Howe, J.Y., 2022. Automatic quantification and classification of microplastics in scanning electron micrographs via deep learning. Sci. Total Environ. 825, 153903. http://dx.doi.org/10.1016/j.scitotenv.2022.153903.

Su, J., Zhang, F., Yu, C., Zhang, Y., Wang, J., Wang, C., Wang, H., Jiang, H., 2023. Machine learning: Next promising trend for microplastics study. J. Environ. Manag. 344, 118756. http://dx.doi.org/10.1016/j.jenvman.2023.118756.

Vattanasit, U., Kongpran, J., Ikeda, A., 2023. Airborne microplastics: A narrative review of potential effects on the human respiratory system. Sci. Total Environ. 904, 166745. http://dx.doi.org/10.1016/j.scitotenv.2023.166745.

Vianello, A., Jensen, R.L., Liu, L., 2019. Simulating human exposure to indoor airborne microplastics using a Breathing Thermal Manikin. Sci. Rep. 9 1, 867. http://dx.doi.org/10.1038/s41598-019-45054-w.

Wegmayr, V., Sahin, A., Saemundsson, B., Buhmann, J., 2020. Instance segmentation for the quantification of microplastic fiber images. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. WACV.

Withana, P.A., Li, J., Senadheera, S.S., Fan, C., Wang, Y., Ok, Y.S., 2024. Machine learning prediction and interpretation of the impact of microplastics on soil properties. Environ. Pollut. 341, 122833. http://dx.doi.org/10.1016/j.envpol.2023.122833.

Wright, S.L., Kelly, F.J., 2017. Plastic and human health: A micro issue? Environ. Sci. Technol. 51 (12), 6634–6647. http://dx.doi.org/10.1021/acs.est.7b00423.

Zarfl, C., Matthies, M., 2010. Are marine plastic particles transport vectors for organic pollutants to the Arctic? Marine Poll. Bull. 60 (10), 1810–1814. http://dx.doi.org/10.1016/j.marpolbul.2010.05.026.

Zhao, H., Wang, X., Yu, X., Peng, S., Hu, J., Deng, M., Ren, L., Zhang, X., Duan, Z., 2024. Application of improved machine learning in large-scale investigation of plastic waste distribution in tourism Intensive artificial coastlines. Environ. Pollut. 356, 124292. http://dx.doi.org/10.1016/j.envpol.2024.124292.

Zhao, X., Zhou, Y., Liang, C., Song, J., Yu, S., Liao, G., Zou, P., Tang, K.H.D., Wu, C., 2023. Airborne microplastics: Occurrence, sources, fate, risks and mitigation. Sci. Total Environ. 858, 159943. http://dx.doi.org/10.1016/j.scitotenv.2022.159943.

Zhen, Y., Wang, L., Sun, H., Liu, C., 2023. Prediction of microplastic abundance in surface water of the ocean and influencing factors based on ensemble learning. Environ. Pollut. 331, 121834. http://dx.doi.org/10.1016/j.envpol.2023.121834.

Zhu, Z., Parker, W., Wong, A., 2023. Leveraging deep learning for automatic recognition of microplastics (MPs) via focal plane array (FPA) micro-FT-IR imaging. Environ. Pollut. 337, 122548. http://dx.doi.org/10.1016/j.envpol.2023.122548.